In the Specification

Please replace the paragraph beginning on page 4 line 8 with the following:

Similarity analysis includes database search and alignment. Examples of public databases include the DNA Database of Japan (DDBJ) ~~(http://www.ddbj.nig.ac.jp/)~~ (www.ddbj.nig.ac.jp/); Genebank ~~(http://www.ncbi.nlm.nih.gov/web/Genbank/Index.htlm)~~ (www.ncbi.nlm.nih.gov/web/Genbank/Index.html); and the European Molecular Biology Laboratory Nucleic Acid Sequence Dabatase (EMBL) ~~(http://www.ebi.ac.uk/ebi_docs/embl_db.html)~~ www.ebi.ac.uk/ebi_docs/embl_db.html). A number of different search algorithms have been developed, one example of which are the suite of programs referred to as BLAST programs. There are five implementations of BLAST, three designed for nucleotide ~~sequences~~ sequence queries (BLASTN, BLASTX, and TBLASTX) and two designed for protein sequence queries (BLASTP and TBLASTN) (Coulson, *Trends in Biotechnology*, *12*:76-80 (1994); Birren, *et al.*, *Genome Analysis*, *1*:543-559 (1997)).

Please replace the paragraph beginning on page 17 line 16 with the following:

Another example of an algorithm that is suitable for determining sequence similarity is the BLAST algorithm, which is described in Altschul et al, J. Mol. Biol. 215: 403-410 (1990). Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information, ~~http://www.ncbi.nlm.nih.gov/~~ www.ncbi.nlm.nih.gov ; *See also* Zhang, Genome Res. 7:649-656 (1997) for the

"PowerBLAST" variation. This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence that either match or satisfy some positive valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul et al, J. Mol. Biol. 215: 403-410 (1990)). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T and X determine the sensitivity and speed of the alignment. The BLAST program uses as defaults a wordlength (W) of 11, the BLOSUM62 scoring matrix (see Henikoff, Proc. Natl. Acad. Sci. USA 89: 10915-10919 (1992)) alignments (B) of 50, expectation (E) of 10, M=5, N=4, and comparison of both strands. The term BLAST refers to the BLAST algorithm which performs a statistical analysis of the similarity between two sequences; see, e.g., Karlin, Proc. Natl. Acad. Sci. USA 90:5873-5787 (1993). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid

to the reference nucleic acid is less than about 0.1, more preferably less than about 0.01, and most preferably less than about 0.001.

Please replace the paragraph beginning on page 48 line 1 with the following:

A microarray-based method for high-throughput monitoring of gene expression may be utilized to measure expression response Schena *et al., Science 270*:467-470 (1995); ~~http://cmgm.stanford.edu/pbrown/array.html~~ cmgm.stanford.edu/pbrown/array.html; Shalon, Ph.D. Thesis, Stanford University (1996). This approach is based on using arrays of DNA targets (e.g. cDNA inserts, colonies, or polymerase chain reaction products) for hybridization to a "complex probe" prepared with RNA extracted from a given cell line or tissue. The probe may be produced by reverse transcription of mRNA or total RNA and labeled with radioactive or flourescent labeling. The probe is complex in that it contains many different sequences in various amounts, corresponding to the numbers of copies of the original mRNA species extracted from the sample.